**Final Project Proposal — MoodTune: Visualizing and Predicting Moods Based on Music Listening Behavior**

**Group Members**: Rob Ranieri, Peter Lin, Gwen Seymour  
**Presentation Date**: June 10, 2025

**1. 🔍 Updated Project Overview**

*MoodTune* is an interactive desktop dashboard that visualizes a user’s music listening behavior in real time and provides mood predictions using machine learning. To comply with Spotify’s Developer Terms, the ML model is trained **offline** using public emotion-labeled music datasets. The pre-trained model is used at runtime to infer mood based on Spotify track features streamed live (not stored). The app offers engaging insights into how music relates to emotional states.

**2. 🎯 Project Objectives**

* Visualize music listening behavior using Spotify audio features
* Predict mood using **pre-trained ML models** (trained on external datasets)
* Provide mood feedback and track suggestions via dashboard
* Enable exploration of genre, time-of-day, and emotional correlations

**3. 🔗 Data Sources**

| **Source** | **Purpose** |
| --- | --- |
| **Spotify Web API** | Real-time audio features (e.g. valence, energy, tempo); no storage used |
| **DEAM Dataset** | External training data: 1,800 music clips labeled with valence/arousal |
| **FMA / Million Song Dataset** | Additional external data for training (optional) |
| **Self-reported moods** | Optional input from users for visualizations (not for training) |

**4. 🔧 Tools & Technologies**

| **Tool** | **Purpose** |
| --- | --- |
| **Python** | Data processing, ML modeling (scikit-learn, pandas) |
| **Spotipy** | Real-time Spotify data ingestion via OAuth |
| **PyQt** | Desktop GUI framework for interactive dashboard |
| **PostgreSQL / Spark** | Store and query training data (not Spotify data) |
| **Seaborn / Plotly** | Visualizations |
| **Jupyter Notebooks** | Model development and experimentation |

**5. 🧠 Machine Learning Strategy**

| **Phase** | **Description** |
| --- | --- |
| **Training** | Train classification (e.g., Random Forest) and clustering (e.g., K-Means) models using only external emotion-labeled datasets (DEAM, FMA). |
| **Model Goals** | Predict emotional categories (happy, sad, angry, relaxed) based on features like valence, energy, tempo. |
| **Inference** | At runtime, stream Spotify audio features and run them through the **pre-trained model** to generate a predicted mood — without storing any Spotify user data. |